The structural basis of protein targeting and translocation in bacteria

Driessen, A.J.M.; Manting, E.H; van der Does, C.

Published in:
Nature Structural Biology

DOI:
10.1038/88549

Publication date:
2001

The structural basis of protein targeting and translocation in bacteria

Arnold J.M. Driessen[1], Erik H. Manting[1,2] and Chris van der Does[1]

In Gram-negative bacteria, two distinct targeting routes assist in the proper localization of secreted and membrane proteins. Signal recognition particle (SRP) mainly targets ribosome-bound nascent membrane proteins, whereas SecB facilitates the targeting of periplasmic and outer membrane proteins. These routes converge at the translocase, a protein-conducting pore in the membrane that consists of the SecYEG complex associated with the peripheral ATPase, SecA. Recent structural studies of the targeting and the translocating components provide insights into how substrates are recognized and suggest a mechanism by which proteins are transported through an aqueous pore in the cytoplasmic membrane.

In Gram-negative bacteria such as *Escherichia coli*, newly synthesized proteins destined either for secretion or incorporation into the membranes are selectively targeted to the translocation machinery at the cytoplasmic membrane. These proteins are synthesized in the cytosol as precursors containing an N-terminal signal sequence (preproteins). Most preproteins are targeted to the cytoplasmic membrane post-translationally *via* the molecular chaperone SecB, whereas many cytoplasmic membrane proteins and some preproteins are targeted cotranslationally to the cytoplasmic membrane by the signal recognition particle (SRP) and its receptor[1,2].

Selection of targeting route is determined immediately after the nascent signal sequence protrudes from the ribosome. SRP binds specifically to long hydrophobic signal sequences and transmembrane segments (for a recent review, see ref. 1). Recognition of less hydrophobic signal sequences by SRP is prevented by the ribosome-associated chaperone trigger factor[3]. At the cytoplasmic membrane, both targeting pathways converge[4] at a protein complex termed 'translocase' (for a recent review, see ref. 5). Translocase is a large, multisubunit membrane-bound enzyme that mediates protein translocation across and into the cytoplasmic membrane. The heterotrimeric SecYEG complex is the central component of the translocase and associates with the cytosolic ATPase SecA, the ribosome[6], YidC[7] and with SecDFyajC[8], another heterotrimeric membrane protein complex. SecA is an ATPase that drives the preprotein translocation reaction and is unique to bacteria. The SecYEG complex is a member of a highly conserved protein translocation pathway[9]. It is homologous to the eukaryotic Sec61p complex, which consists of three subunits (α, β and γ) that together form the 'translocon' of the endoplasmic reticulum (ER) membrane[10]. YidC is a membrane protein that is homologous to Oxa1p, a mitochondrial protein that acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA[11]. YidC works either independently or together with the SecYEG complex to insert proteins into the cytoplasmic membrane, depending on the protein being inserted[7,12].

Major unresolved questions in protein translocation relate to how the cytosolic targeting components select their substrates and how they cooperate to deliver the protein substrates to the translocation machinery. Structural analyses of the targeting components provide insights into how different signal peptides and protein substrates are efficiently recognized despite considerable sequence variation. Moreover, these analyses may identify sites of interactions between the targeting components, thereby providing insights into the mechanism of substrate delivery. Other questions concern the energetics of protein translocation, in particular how ATP and the proton motive force are utilized to drive unfolded preproteins across the membrane. Studies on the nucleotide-binding components will provide clues about how nucleotide binding and hydrolysis are coupled to translocation. It is also not known how the translocase can catalyze such diverse functions as the transmembrane passage of preproteins through an aqueous pore in the membrane or the insertion of hydrophobic transmembrane segments of proteins into the phospholipids bilayer. For these distinct functions, the SecYEG complex may transiently associate with other integral membrane proteins such as SecDFyajC, YidC or other yet unidentified factors. Understanding the structure and dynamics of the translocase is, therefore, essential for studying the mechanism of the translocation reaction.

Here we review recent crystallographic and electron microscopy studies of the components of the bacterial targeting and translocation pathway that have led to a remarkable progress in our understanding how these systems operate and how proteins cross the membrane. The review focuses on the targeting components SRP and SecB, and how they interact with the protein substrates and other ligands. We also discuss recent structural studies that provide insights into the mechanism by which SecYEG assembles into a protein-conducting channel.

Structural basis of SRP-mediated protein targeting

SRP-mediated protein targeting has been studied in great detail in the ER of mammals[13]. The mammalian SRP is a ribonucleoprotein that consists of six proteins and a 7S RNA molecule that is 300-base pairs long. The 54 kDa polypeptide subunit (SRP54) is a GTPase and recognizes the hydrophobic signal sequence at the N-terminus of the nascent chain. SRP targets the ribosome-nascent chain complex (RNC) to the ER membrane, where the SRP54 subunit binds to another GTPase, SRα, the peripheral subunit of the membrane-bound heterodimeric SRP receptor. This interaction stimulates GTP binding to both proteins and promotes the release of the nascent chain to the protein-conducting channel in the ER

[1]Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands.
[2]Present address: Department of Clinical Viro-immunology, Sanquin-CLB, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands.

Correspondence should be addressed to A.J.M.D. email: *a.j.m.driessen@biol.rug.nl*

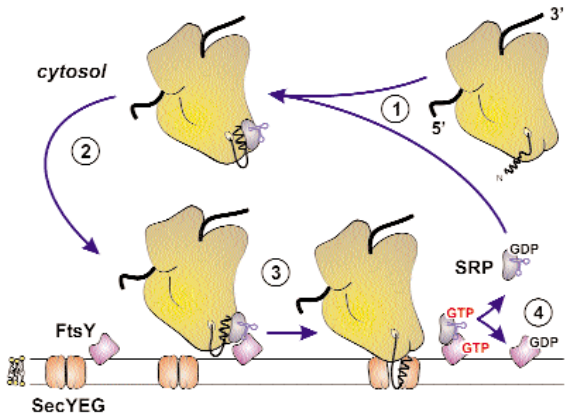

Fig. 1 Model for SRP-mediated targeting of ribosome-nascent chain complexes. For description see text.

membrane. GTP hydrolysis then facilitates the recycling of SRP into the cytosol.

SRP-mediated protein targeting in bacteria. Bacteria have a similar, perhaps less complicated SRP system[1]. In *E. coli*, the SRP complex is composed of only a 48 kDa GTPase termed Ffh (for fifty-four-homolog) that exhibits marked homology throughout its sequence with the mammalian SRP54 and a ~100-base pair long 4.5 S RNA that is nearly identical in structure to the SRP54 binding domain of the eukaryotic SRP RNA (domain IV). FtsY is the bacterial homolog of SRα. Notably, all the bacterial SRP components are essential for cell growth.

Synthesis of secretory and membrane protein begins with an unattached ribosome in the cytosol (Fig. 1). In the SRP-mediated targeting pathway, the SRP complex recognizes the signal sequence or a hydrophobic transmembrane segment emerging from the ribosome (Fig. 1, step 1). The RNC-bound SRP is then targeted to FtsY, the membrane-bound receptor; this interaction increases the GTP binding affinity of Ffh and FtsY (Fig. 1, step 2)[14]. The 4.5S RNA increases the affinity of Ffh for signal sequences[15]; it is also essential for the interaction between SRP and its receptor, FtsY[16]. In bacteria, there is no known counterpart to the eukaryotic SRβ, an integral membrane protein that serves to anchor SRα to the membrane, and FtsY is distributed between the cytosol and the membrane. Binding of GTP to the membrane-bound FtsY and/or SRP dissociates the RNC from SRP and releases it to the SecYEG complex (Fig. 1, step 3). Crosslinking studies provided evidence that these two processes are linked[4], whereas the large ribosomal subunit has been shown to bind with high affinity to the SecYEG complex *via* the 23S rRNA[6]. Hydrolysis of GTP dissociates SRP from FtsY, allowing SRP to recycle into the cytosol while FtsY remains bound to the membrane (Fig. 1, step 4). The polypeptide chain then continues to elongate until translation is complete.

Although SecA is not needed during the SRP-dependent targeting and the initial insertion into the translocase, it is required for the overall translocation reaction. Therefore, it is questionable whether the ribosome alone is sufficient to drive SecYEG-mediated preprotein translocation in bacteria. For instance, during the insertion of proteins with several transmembrane segments (polytopic transmembrane proteins), the ribosome and SecA collaborate[17] — in this process, SecA is needed to translocate hydrophilic polypeptide segments across the membrane[17]. Although crosslinking studies with nascent membrane proteins show that RNCs contact both SecY and SecA[4], it is not clear whether the ribosome and SecA bind SecYEG simultaneously.

The exact steps that initiate GTP binding and hydrolysis have not yet been precisely defined in bacteria, nor is it fully understood how Ffh and FtsY interact, nor how FtsY associates with the membrane. The ease of purification and perhaps the reduced complexity of the bacterial SRP components have in recent years enabled structure determination of domains of SRP[18–21] and FtsY[22] (Fig. 2). A focal point of these structural studies has been the detailed understanding of how different hydrophobic signal peptides are recognized efficiently, why conserved residues in Ffh and RNA are required for SRP assembly and how nucleotide binding affects the conformation of the protein domains.

Structures of Ffh and FtsY. Ffh contains three structural domains: an N-terminal four-helix bundle (N-domain), a central Ras-like GTPase domain (G-domain) and a methionine-rich C-terminal domain (M-domain)[19]. FtsY also consists of three domains: a strongly acidic N-terminal domain, a central N-domain and a C-terminal GTPase domain (G-domain). The NG-domains of Ffh and FtsY are structurally related (compare Fig. 2*a*,*d*)[18,22].

In Ffh and FtsY, the N-domain is closely associated with the adjacent G-domain, which contains the GTP binding site with four characteristic sequence motifs (I–IV)[18,20,22]. A comparison of the structures of the Mg^{2+}-GDP-bound and apo forms of the *Thermus aquaticus* Ffh NG-domain suggests a mechanism of the GTPase function[20]. Mg^{2+} is required for nucleotide binding and catalysis. In the absence of Mg^{2+}-GDP, the conserved side chains of the empty nucleotide binding site (motifs I–III) form a tight network of interactions that may stabilize this nucleotide-free form. In the Mg^{2+}-GDP bound form, motifs I–III interact with the Mg^{2+} and β-phosphate, while motif IV recognizes the guanine base (Fig. 2*c*)[20]. Therefore, a substantial structural rearrangement must accompany the formation of the catalytically active complex of Ffh with Mg^{2+}-GTP. It is difficult to predict the exact conformational change induced by GTP hydrolysis without the structure of the Ffh–Mg^{2+}-GTP complex. However, it has been suggested that upon hydrolysis of the γ-phosphate bond, the β-phosphate of GDP rotates to bind a conserved glutamine (position 144 in Ffh of *T. aquaticus*). The altered interaction of the β-phosphate and the release of the γ-phosphate are thought to open a loop covering the nucleotide binding site that is specific to the SRP GTPase and induce a movement of motif IV relative to the N-domain[20]. This hypothesis can now be tested by site-directed mutagenesis of the critical residues involved.

The interface between the N- and G-domains harbors many conserved residues that are important for function. By coupling the position of the N-domain relative to the position of motif IV of the G-domain (Fig. 2*c*), the nucleotide occupancy of the G-domain can either be sensed or controlled. This conformational change may communicate the catalytically active GTP-bound state between the SRP and its receptor FtsY, but the exact contact interface between Ffh and FtsY is unknown. Although the different conformers described above are also observed in the apo form of Ffh NG domain[19,20] — apparently trapped by crystal packing forces — it is believed that these different nucleotide-bound forms are real conformational intermediates. GTPases usually exhibit a very high affinity for GDP and, therefore, require a nucleotide release factor to mediate GDP-GTP exchange. The SRP GTPases, however, exhibit a unique mechanism of GDP release and stabilization of the empty form. This may explain their low *in vitro* GDP binding affinity and the absence of a separate nucleotide release factor[20,23].

Role of SRP RNA. The M-domain of Ffh binds both SRP RNA and signal peptides[13]. The crystal structure of the *T. aquaticus* Ffh M-domain shows a deep groove that comprises almost exclusively

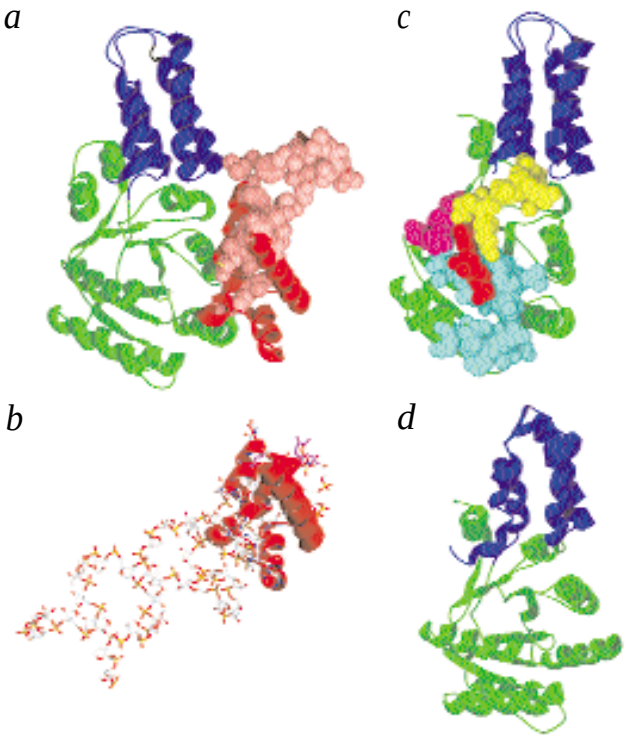

Fig. 2 Crystal structures of SRP pathway components. ***a***, Ribbons drawing of Ffh of *T. aquaticus* (PDB accession no: 2FFH)[19]**.** The domains are indicated with different colors: N-domain, blue; G-domain, green; and M-domain, red. Amino acid residues of the M domain implicated in signal sequence binding are indicated as space-filling models in pale pink. ***b***, Ribbons drawing of the SRP ribonucleoprotein core of *E. coli* (1HQ1)[21] showing the M-domain with domain IV of the 4.5S RNA. The RNA is shown as a stick model, and the drawing is in the same orientation as in (*a*). ***c***, Ribbons drawing of the NG domain of Ffh of *T. aquaticus* (2NG1)[20] with bound Mg^{2+}-GDP. The color scheme is the same as in Fig. 2*a*. GDP is indicated in red. The amino acid residues in motifs I–III in the GTPase domain are indicated in pale blue, and those in motif IV are indicated by yellow. The amino acid residues in the closure loop specific to SRP GTPase are shown in purple. ***d***, Ribbons drawing of the NG domain of the *E. coli* FtsY (1FTS)[22]. The domains are indicated in the same color as in Fig. 2*a*. This fragment of FtsY lacks the N-terminal acidic domain.

the side chains of hydrophobic residues (Fig. 2*a*)[19]. The size of the groove and its conserved hydrophobic character suggest that it forms the signal sequence binding pocket of SRP. This region contains the so-called 'methionine-bristle', which has been implicated in signal sequence binding[13]. The structure of the *E. coli* Ffh M-domain associated with a fragment of SRP RNA corresponding to the conserved domain IV shows some surprising and unusual features (Fig. 2*b*)[21]. First, the M-domain recognizes a distorted RNA minor groove. Second, part of the RNA lies adjacent to the groove that has been implicated in binding signal sequences. This suggests that the signal sequence binding site may consist of both protein and RNA. Batey *et al.*[21] proposed that signal sequences bind to the M-domain–RNA complex *via* a combination of hydrophobic and electrostatic interactions. Ultimate proof for such an interaction, however, requires a structure of the M-domain–RNA complex with a bound signal peptide. Since binding of 4.5S RNA to Ffh also influences the kinetics of the Ffh–FtsY interaction[16], one might speculate that binding of the signal sequence to the M-domain of Ffh results in small changes in the RNA structure that alters the Ffh–FtsY interaction, which enables binding of GTP to both proteins. To substantiate this hypothesis, the structure of Ffh–FtsY complex will also be required.

Structural basis of SecB-mediated protein targeting

SecB is a molecular chaperone required for efficient protein export in most Gram-negative bacteria. It is a highly acidic homotetrameric protein with a subunit molecular mass of ~17 kDa (for a recent review see ref. 2). SecB does not interact with the signal sequence and recognizes longer nascent chains than SRP[24]. The SecB-bound preproteins are then targeted to the SecYEG-bound SecA protein[25]. In this targeting pathway (Fig.3), SecB binds to the mature domain of a nascent preprotein (Fig. 3, step 1), and stabilizes it in denatured conformations (Fig. 3, step 2)[26]. These conformations, which most likely contain native-like secondary structure elements without specific tertiary interactions, are collectively referred to as the 'translocation competent' state of the preprotein. Stably folded preproteins are not translocated by the translocase. The SecB–preprotein complex is then targeted to the SecYEG-bound SecA (Fig. 3, step 3), but it may first associate with low affinity binding site with cytosolic SecA and remain in the cytosol until translocation sites at the membrane become available. Targeting to the translocase requires the high affinity binding of SecB to the C-terminus of SecYEG-bound SecA[22]. The signal sequence of the preprotein then binds to SecA, and this tightens the SecB–SecA interaction, the site of which is unknown, and induces the release of the preprotein from SecB with the concomitant transfer to SecA (Fig. 3, step 5)[27]. The release of SecB from the membrane is coupled to the binding of ATP to SecA (Fig. 3, step 6)[27].

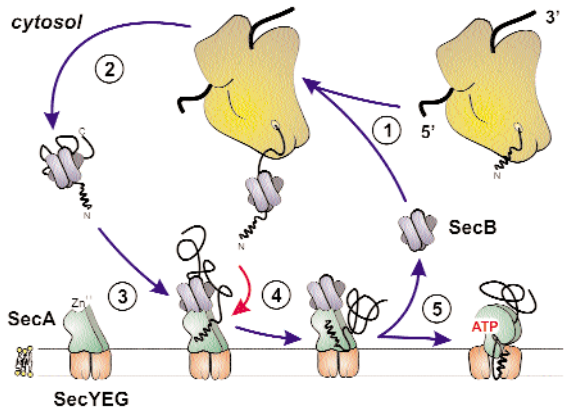

Fig. 3 Model for SecB-mediated protein targeting. For description see text.

For a long time it was unclear how SecB simultaneously recognizes both ligands — the denatured preprotein and SecA — by entirely different types of interaction. The crystal structure of the *Haemophilus influenzae* SecB[28] solved recently provides new insights into these activities of SecB.

SecB structure. The SecB monomer has a simple $\alpha + \beta$ fold (Fig. 4*a*)[28]. The tetrameric protein is organized as a dimer of dimers, which is consistent with the observed dynamic dimer-tetramer equilibrium of SecB in solution[29]. SecB has been crystallized without a peptide substrate in its binding site, but the structure reveals the presence of four grooves located at the sides of the tetramer. Two grooves at the same side of the molecule form a 70 Å long surface-exposed channel — and thus two channels per tetramer — that crosses the dimer–dimer interface (Fig. 4*c*). This channel is equipped with all the characteristics needed to bind a wide range of peptide substrates.

By screening an extensive peptide library for SecB binding activity, it has been possible to define a SecB binding motif[30]. Typical peptide substrates are ~9 residues long and enriched in aromatic

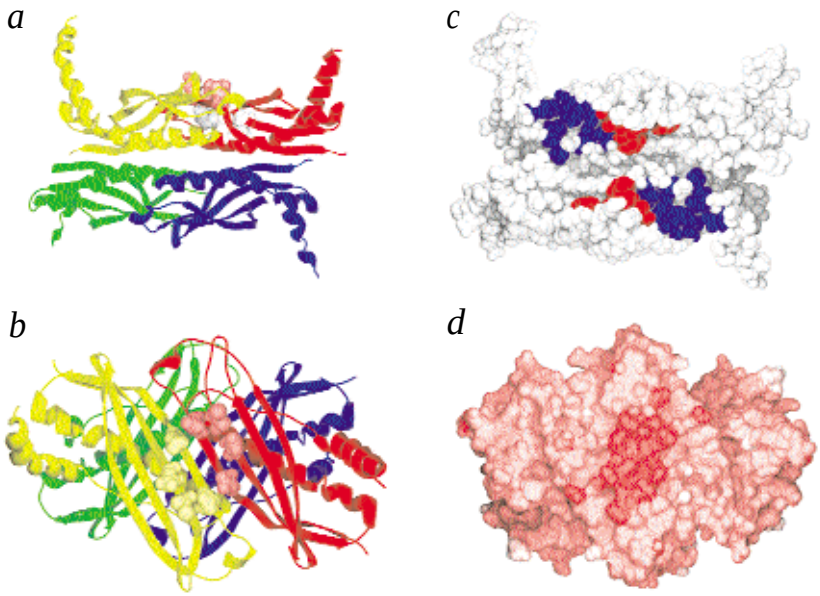

Fig. 4 Crystal structure of the *H. Influenza* SecB protein (1FX3)[28]. ***a***, Ribbon drawing of the side view of the SecB tetramer. The four subunits are shown in red, yellow, blue and green. The amino acids residues of the *H. influenzae* SecB (Asp27, Glu31, Glu86 and Ile84) that match those residues of *E. coli* SecB that have been implicated in SecA binding are colored in pale red. Corresponding residues of the *H. influenzae* SecB (Phe 83, Cys 85, Val 87 and Gln 89) that, when mutated in *E. coli* SecB, disrupt the binding of preproteins are shown in white. These mutations also shift the dimer-tetramer equilibrium towards the dimer. ***b***, Ribbon drawing of the top view of the SecB tetramer. The color of the subunits is as in (*a*). The positions of the residues implicated in SecA binding are highlighted on two subunits (in pale red and yellow), showing their clustering on the surface. ***c***, The proposed peptide binding channel. The solvent-accessible surface is shown in the same orientation of the molecule as in Fig. 4*a*. The proposed subsite 1 and 2 are highlighted in red and blue, respectively. ***d***, The proposed SecA binding site. The solvent accessible surface is colored based on the electrostatic potential (ranging from -25 to +25 kT). The molecule is shown in the same orientation as in Fig. 4*b*.

and basic residues; acidic residues are strongly disfavored. The majority of the binding energy for these peptides results from hydrophobic interactions[31]. Xu *et al.*[28] proposed that each peptide binding groove can be roughly subdivided in two subsites that may recognize distinct features of the preprotein (Fig. 4*c*). Subsite 1 is a deep cleft located in a narrow constriction in the middle of the long groove. Most of the amino acids that line its surface are conserved aromatic residues, which are ideal for the binding of hydrophobic and aromatic regions of polypeptides. The conformational fluctuation among the different subunits indicates that subsite 1 is structurally flexible, providing the plasticity needed to accommodate the wide range of peptides that bind to SecB. Subsite 2 is a shallow, open groove with a hydrophobic floor. This site is devoid of aromatic residues. It may bind to β-pleated segments that are abundantly present in the preprotein substrates by forming regular main chain hydrogen bonds. Since SecB has been crystallized without a peptide substrate in its binding groove, it cannot be excluded that subsite 2 also binds other structural elements. Negatively charged residues positioned at the rim of the long channel could provide the selectivity for basic residues.

SecB forms a stoichiometric complex with preprotein substrates that typically have a molecular mass of 30–50 kDa. Carboxy methylated bovine pancreatic trypsin inhibitor, a small stably unfolded model substrate, binds at four distinct sites on SecB that show little cooperativity[2]. These sites likely correspond to the four peptide binding grooves identified in the SecB structure. Mutations in the *E. coli* SecB (C76Y, V78F and Q80R) that disrupt the interaction between SecB and preproteins shift the tetramer-dimer equilibrium towards the dimer[29]. These residues do not map in the peptide binding grooves but are localized on a surface-exposed β-strand where they point towards the interior of the molecule (Fig. 4*a*). Although they do not participate in the dimer–dimer interface, they probably indirectly disrupt this interface and thereby impair preprotein binding at the long peptide binding channels that overlap with the dimer–dimer interface.

Substrate recognition. How does SecB differentiate between secreted and cytosolic proteins? The signal sequence does not bind to SecB with detectable affinity; rather, it retards the folding of the mature region of a protein[26]. SecB recognizes peptide segments that are typically found within the interior of folded proteins[30], and thus preferentially binds to the unfolded conformation of the mature part of preproteins[32]. SecB associates with ribosome-bound nascent chains after they have reached a length of ~150 residues[24]. These long binding regions have been proposed to simultaneously occupy multiple binding sites on SecB, thereby allowing a high affinity of interaction (K_d 5–50 nM)[33]. To access the peptide binding channels on both sites, the long unstructured polypeptide segments are presumably wrapped around the chaperone. In this manner, SecB stabilizes the preprotein in a translocation-competent state.

SecA recognition. SecB shares the ability to stabilize the unfolded state of preproteins with general chaperones such as GroEL and DnaK. These general chaperones, however, fail to stimulate translocation, as they are unable to target the preprotein to the translocase correctly[2]. SecB binds the SecYEG-associated SecA with high affinity (K_d 10–30 nM)[25]. The C-terminus of SecA serves as a specific binding site for SecB[27]. This region is highly conserved among most bacterial SecA proteins, is positively charged and bears a zinc ion that is needed for the functional interaction between SecB and SecA[2]. The well-conserved amino acid residues, Asp 20, Glu 24, Leu 75 and Glu 77, of the *E. coli* SecB are important for the high affinity interaction with SecA[34]. A remarkable feature of the *H. influenza* SecB structure is that the corresponding residues cluster on a flat, solvent-exposed acidic surface that is present on both sides of the molecule (Fig. 4*d*)[28]. This surface is ideally suited to interact electrostatically with the positively charged SecB-binding domain on SecA. The C-terminus of SecB has also been implicated in SecA binding. This highly flexible, solvent-exposed region is barely resolved in the SecB structure, but seems to protrude from the structure as long 'arms' (Fig. 4*a*,*c*). These arms possibly embrace the SecB binding site on SecA to facilitate subsequent preprotein transfer.

The symmetry in the SecB structure provides a nice solution to the dilemma of how the SecB tetramer binds the homodimeric SecA[27]. Each subunit of the SecA dimer could interact with the negatively charged surface formed by one dimer of SecB. To fully appreciate how this relates to the positions of the C-termini of the SecA dimer, the long awaited structure of SecA is needed.

SecA modulates the preprotein binding affinity of SecB, but the structural basis of this phenomenon is not known. Some of the mutations in SecB that interfere with SecA and preprotein binding map to different faces of a surface exposed β-strand (see ref. 23 and

refs therein) (Fig. 4*a*). One might speculate that the binding of SecA to SecB causes a conformational change that is sensed by amino acid residues on the opposite face of the β-strand, and that propagates to the long α-helix at the SecB dimer–dimer contact interface. This would result in a destabilization of the long peptide-binding channel and a reduction of the preprotein binding affinity.

The structural and biochemical studies indicate that SecB uses its entire molecular surface to interact with preprotein and SecA. For preprotein transfer and its own release from the membrane, SecB relies on the catalytic activity of SecA[27] (Fig. 3). However, to understand exactly how SecA accepts the preprotein from SecB, and how it releases SecB from the membrane, the structure of the ternary preprotein–SecB–SecA complex is required. The location and characteristics of the peptide binding subsites can now be directly tested by site-directed mutagenesis.

The preprotein-conducting channel

Translocase can be dissected into two modules, consisting of a protein-conducting pore formed by the transmembrane SecYEG complex and a unit that directs movement of the translocating polypeptide chain. For post-translational translocation of preproteins, SecYEG associates with the motor protein SecA[5], while for cotranslational membrane protein insertion, it can directly bind to the ribosome[6].

Experimental evidence shows that translocase comprises the combined properties of a translocation pore and a molecular motor. First, the addition of translocation ligands to isolated *E. coli* membranes opens a proton-conducting channel, and the conductance is correlated to the expression levels of SecYE complex[35]. Second, a translocating preprotein is in close proximity to the SecA and SecY proteins[36]. Third, SecYEG-bound SecA is shielded from phospholipid[37]. It thus appears as if the translocase is involved in creating a protein-conducting pore across the membrane that is lined by protein, not phospholipid. Fourth, SecA functions as an ATP-driven molecular motor. It binds the preprotein, and in an ATP-dependent manner, threads it in an unfolded state through the SecYEG complex[38].

ATP is the essential energy source for protein translocation; however, the proton motive force also stimulates the rate of translocation. Translocation is initiated by the binding of ATP to SecA, which drives the membrane insertion of a loop of the signal sequence and early mature domain to an extent that the signal sequence can be processed by signal peptidase (the catalytic site is exposed to the periplasm)[38]. Upon ATP hydrolysis, SecA releases the bound preprotein to the SecYEG complex. Further translocating steps can be driven by ATP and/or the proton motive force. In the absence of proton motive force, the catalytic cycle of the SecA ATPase can be divided into two separate translocation steps[39]. SecA can rebind to the partially translocated polypeptide chain, and this event causes the translocation of a stretch of ~20 amino acids. Another ~20 amino acids are translocated upon the binding of ATP to SecA. Repeated cycles of ATP binding and hydrolysis, as well as preprotein binding and release permit the stepwise translocation of the preprotein across the membrane[38–40]. Once SecA has released the preprotein upon ATP hydrolysis, the proton motive force can drive the translocation reaction by a mechanism that has remained unsolved. This is one of the main challenges for future studies.

A central question is how the SecYEG complex can encompass a channel that conducts the transmembrane movement of polypeptides. Does such a channel consist of a single SecYEG complex possibly gated by SecA and preprotein, or do multiple SecYEG complexes assemble into a pore-like structure? Is such a pore a passive device or does it actively participate in translocation? Halting the SecA-dependent translocation reaction by addition of the nonhydrolyzable ATP analog AMP-PNP results in the crosslinking of two neighboring SecE molecules[41]. When assuming a heterotrimeric stoichiometry of SecYEG, this experiment indicates the presence of more than one SecYEG complex per translocase. This postulate is confirmed by recent structural analyses of the SecYEG complex.

SecYEG forms an oligomeric protein-conducting channel. A first glimpse of the possible structure of SecYEG was revealed by structural analyses of eukaryotic translocon, the Sec61p complex[42]. The α-and γ-subunit of Sec61p are homologous to SecY and SecE, respectively, but the β-subunit and SecG are distinct. The reported shape of the ribosome-bound purified Sec61p complex visualized by cryo electron microscopy (cryo-EM) is a roughly pentagonal ring, 5–6 nm high and ~9 nm wide. It has a cylindrical pore of ~2 nm in diameter that extends throughout the protein complex perpendicular to the plane of the membrane[43]. Ribosome–Sec61p complexes purified from membrane vesicles derived from the ER (microsomes) contain in addition the translocon associated protein (TRAP) and an oligosaccharyl transferase, and show a larger elliptical pore with a size of 2 × 5 nm (ref. 44). The formation of channels with purified and reconstituted Sec61p required prior incubation with ribosomes or co-reconstitution with the yeast Sec62/63 accessory proteins involved in post-translational translocation[42]. The absence of the channel structures with reconstituted Sec61p and their reassembly upon interaction with translocation ligands may represent naturally occurring assembled and disassembled states. Alternatively, a lowered stability of the Sec61p protein in detergent solution may result in the dissociation of the complexes during purification. With respect to the latter possibility, the purified *Bacillus subtilis* SecYE[45] appears more stable as it gave rise to ring-like structures without the requirement of prior incubation with translocation ligands.

The suggestion that the central pore in Sec61p could be part of a protein-conducting channel has been supported by the three-dimensional reconstruction of the ribosome–Sec61p complex structures from EM images[43]. The exit channel of the large ribosomal subunit aligns with the putative protein-conducting pore of the Sec61p complex to form a continuous protein conduit. The presence of a 1.5–2 nm pore formed by the Sec61p complex is consistent with conductivity experiments using microsomal membranes under nontranslocating conditions[46]. However, under translocation conditions, pores with an estimated size of ~5 nm have been deduced from ion fluxes and the size of molecules that can be used to quench fluorescent probes attached to translocating nascent polypeptide chains[46]. Also in bacteria, a pathway that allows passage of large substrates across the membrane is expected, as preproteins with a stable disulfide bonded tertiary loop of a maximum of 20 amino acids can be transported across the membrane provided that ATP and the proton motive force are present as energy sources[47].

A first image of the integral membrane part of the bacterial translocase has been obtained by EM studies of *B. subtilis* SecYE. SecG is apparently not required for the basic structural organization of the SecYEG complex, as structures of 7.0–8.5 nm in diameter resembling the pentagonal-like Sec61p structure were observed[45]. Negative-stain EM images of solubilized *E. coli* SecYEG revealed protein structures with dimensions similar to those reported for the *B. subtilis* SecYE but with an apparent two-fold symmetry[48] (Fig. 5*a*). In addition, smaller particles were observed that possibly represent monomeric SecYEG complexes[48]. The protein mass calculated from the *B. subtilis* particle dimensions sug-

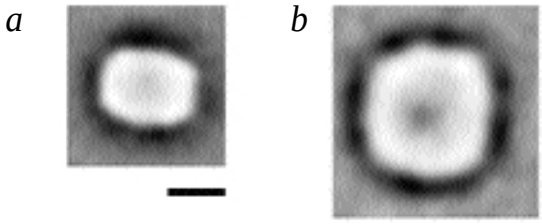

Fig. 5 Low resolution projection maps of the SecYEG structures[48]. ***a***, Dimeric SecYEG complexes have dimensions of 6.7 × 8.7 nm with a central 2 nm wide stain-filled indentation. ***b***, Tetrameric SecYEG complexes have a square-like shape and a size of 10–12 nm with a 4–6 nm central stain-filled cavity. The bar denotes 5 nm. The major class averages were generated by single particle analysis after electron microscopy of negatively stained SecYEG complexes solubilized in dodecylmaltoside. The large SecYEG structures were obtained after incubation with SecA and AMP-PNP and subsequent purification. Another class of particles represented by monomeric SecYEG is not shown here.

gests an assembly of three (SecYE) complexes[45]. However, mass measurements using scanning transmission electron microscopy (STEM) indicated the presence of two SecYEG complexes in the corresponding *E. coli* particles[48]. A possible explanation for this apparent discrepancy could be the presence of varying amounts of phospholipid and/or detergent bound to the solubilized complexes.

When the reconstituted *E. coli* SecYEG complex was incubated with SecA and the nonhydrolyzable ATP analog AMP-PNP, purified resolubilized SecYEG contained substantially larger particles[48]. These particles encompass a 5 nm wide stain-filled central pore or indentation and have an outer diameter of 10.5–12 nm (Fig. 5*b*). STEM analysis of the larger particles indicated the presence of four SecYEG complexes. This number of SecYEG complexes surrounding the active protein translocation channel was consistent with the mass of detergent-soluble translocase stabilized by a preprotein translocation intermediate and the SecA:SecY stoichiometry (1:2) in this soluble translocase–preprotein complex[48].

Translocation of preproteins is initiated by the binding of ATP to the SecYEG- and preprotein-bound SecA. *In vitro*, ATP can be replaced by AMP-PNP to initiate translocation, but the process is then halted at that stage. It thus appears that the nucleotide-induced SecA-dependent initiation of translocation creates a large translocation pore across the membrane that is formed by four SecYEG complexes. This structure is recruited from multiple copies of monomeric or dimeric SecYEG and is only stable under translocation conditions[48].

SecA-mediated preprotein-translocation

The driving force for the unidirectional movement of the translocating polypeptide chain *via* SecYEG is generated on the cytosolic-side (pushing) by the SecA ATPase. How does SecA mediate the stepwise translocation of preprotein segments across the membrane? SecA functions as a dimer, and each monomer consists of at least two domains, a C-terminal 34 kDa domain and an N-terminal ATPase domain of 68 kDa (for a recent review see ref. 5). In the absence of translocation ligands, the ATPase activity of the N-terminal domain is down regulated by an interaction with the C-terminal domain[49,50]. The SecYEG-bound SecA is activated for ATPase activity when it associates with a preprotein. Under these conditions, membrane-protected SecA proteolytic fragments can be generated that correspond to the N- and C-terminal domains[40,51]. Other studies show that some portions of the SecA surface are accessible for chemical labeling and proteolysis from the periplasmic side of the membrane[52]. This has been taken to imply that the entire SecA protein penetrates the membrane when it is activated by ATP and preproteins[51]. According to the 'membrane-insertion' hypothesis, preprotein segments are translocated across the membrane by coinsertion with SecA[40].

The protease-resistant conformation of SecA correlates with an active state of the SecA protein during translocation. This has been interpreted as the ATP-bound, membrane-inserted conformation of SecA[40,51]. However, experimental and conceptual arguments contradict the membrane insertion hypothesis. The SecA fragments described above can also be generated in the absence of a membrane environment when a detergent is used that preserves the SecA–SecYEG interaction[53]. Small angle X-ray scattering experiments indicate that the SecA dimer is an elongated molecule with a maximal width of 8 nm and length of 15 nm (ref. 54). The thickness of the lipid bilayer is insufficient to accommodate the entire SecA molecule or large domains of it. The recent proposal that SecYEG functions as a monomer[55] also seems inconsistent with the membrane insertion hypothesis of SecA function as it is unclear how the monomeric SecYEG complex with dimensions of ~4.8 × 6.5 nm (ref. 48) can accommodate the entire SecA structure without considerable rearrangements of the packing of the transmembrane segments of the SecYEG complex[41]. Also, the possibility exists that chemical agents and proteinase added from the periplasmic side of the membrane gain access to SecA *via* the pore formed by the tetrameric SecYEG complex[48]. Since conformational changes in SecA that lead to proteinase protection may not represent membrane insertion, other mechanisms that do not involve the insertion of large domains should be considered. For instance, insertion of only the SecA preprotein-binding loop into the protein-conducting channel may suffice to push polypeptide segments across the membrane. The availability of the SecA structure will stimulate new biochemical experiments that will address the possible displacement of the preprotein-binding loop of SecA by, for instance, site-specific crosslinking or fluorescence techniques.

Perspectives

The progress that has been achieved in recent years to understand the functional and structural features of the bacterial targeting and translocation components is impressive. One of the challenges that lies ahead is to understand the structural basis of how the components cooperate in targeting, and how they transfer (nascent) preproteins to the translocase. This probably requires structures of the intact SRP components as a complex, both in the presence and absence of peptide and nucleotide ligands. The SecB structure provides a first glimpse of how the ligands may bind, but to fully understand its substrate specificity and the downstream transfer mechanism, structures of SecB with bound polypeptide substrate in combination with the SecA C-terminus or the entire SecA protein are needed. The high resolution structure of SecA will be a step towards reaching this goal. This structure alone will be of great interest and will tell its own story about the complicated ATPase coupling mechanism. The various models for SecA-mediated translocation will likely stir controversy until the surface topography of the SecYEG-bound SecA in the presence of nucleotide ligands is determined. The ultimate proof that SecA passes proteins through an aqueous channel requires a direct observation of a preprotein moving through the center of the channel-like structure.

During the SecYEG-mediated insertion of proteins into the membrane, hydrophobic transmembrane segments should be able to leave the translocation channel laterally. It will be very interesting to obtain structural information on the translocation channel during this lateral opening process. This will reveal a possible

structural difference in channel structure during protein translocation and membrane protein integration. YidC could have a role in this process as it associates with the hydrophobic transmembrane segments[7]. Current structures of the SecYEG complex are of low resolution. For a detailed understanding of the mechanism of protein translocation, high resolution structures of the different states of this fascinating protein complex or even the complete translocase are required.

Acknowledgments

We would like to thank A. Mark and N. Nouwen for careful reading of the manuscript, and the past and present members of the team who contributed to this work. This work was supported by the Earth and Life Sciences Foundation (Algemene Levenswetenschappen, ALW), the Netherlands Foundation for Chemical Research (Chemische Wetenschappen, CW), and the Netherlands Foundation for Scientific Research (Nederlandse Wetenschappelijk Onderzoek, NWO).

Received 9 January, 2001; accepted 18 April, 2001

1. Herskovits, A.A., Bochkareva, E.S. & Bibi, E. New prospects in studying the bacterial signal recognition particle pathway. *Mol. Microbiol.* **38**, 927–939 (2000).
2. Fekkes, P. & Driessen, A.J.M. Protein targeting to the bacterial cytoplasmic membrane. *Microbiol. Mol. Biol. Rev.* **63**, 161–173 (1999).
3. Beck, K., Wu, L.F., Brunner, J. & Müller, M. Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J.* **19**, 134–143 (2000).
4. Valent, Q.A. *et al*. The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. *EMBO J.* **17**, 2504–2512 (1998).
5. Manting, E.H. & Driessen, A.J.M. *Escherichia coli* translocase: the unraveling of a molecular machine. *Mol. Microbiol.* **37**, 226–238 (2000).
6. Prinz, A., Behrens, C., Rapoport, T.A., Hartmann, E. & Kalies, K.U. Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. *EMBO J.* **19**, 1900–1906 (2000).
7. Scotti, P.A. *et al*. YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J.* **19**, 542–549 (2000).
8. Duong, F. & Wickner, W. The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. *EMBO J.* **16**, 4871–4879 (1997).
9. Hartmann, E. *et al.* Evolutionary conservation of components of the protein translocation complex. *Nature* **367**, 654–657 (1994).
10. Johnson, A.E. & van Waes, M.A. The translocon: a dynamic gateway at the ER membrane. *Annu. Rev. Cell Dev. Biol.* **15**, 799–842 (1999).
11. Hell, K., Neupert, W. & Stuart, R.A. Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *EMBO J.* **20**, 1281–1288 (2001).
12. Samuelson, J.C. *et al.* YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637–641 (2000).
13. Walter, P. & Johnson, A.E. Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10**, 87–119 (1994).
14. Powers, T. & Walter, P. Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases. *Science* **269**, 1422–1424 (1995).
15. Zheng, N. & Gierasch, L.M. Domain interactions in *E. coli* SRP: stabilization of M domain by RNA is required for effective signal sequence modulation of NG domain. *Mol. Cell* **1**, 79–87 (1997).
16. Peluso, P. *et al*. Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor. *Science* **288**, 1640–1643 (2000).
17. Neumann-Haefelin, C., Schafer, U., Muller, M. & Koch, H.G. SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. *EMBO J.* **19**, 6419–6426 (2000).
18. Freymann, D.M., Keenan, R.J., Stroud, R.M. & Walter, P. Structure of the conserved GTPase domain of the signal recognition particle. *Nature* **385**, 361–364 (1997).
19. Keenan, R.J., Freymann, D.M., Walter, P. & Stroud, R.M. Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell* **94**, 181–191 (1998).
20. Freymann, D.M., Keenan, R.J., Stroud, R.M. & Walter, P. Functional changes in the structure of the SRP GTPase on binding GDP and Mg^{2+}GDP. *Nature. Struct. Biol.* **6**, 793–801 (1999).
21. Batey, R.T., Rambo, R.P., Lucast, L., Rha, B. & Doudna, J.A. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* **287**, 1232–1239 (2000).
22. Montoya, G., Svensson, C., Luirink, J. & Sinning, I. Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. *Nature* **385**, 365–368 (1997).
23. Moser, C., Mol, O., Goody, R.S. & Sinning, I. The signal recognition particle receptor of *Escherichia coli* (FtsY) has a nucleotide exchange factor built into the GTPase domain. *Proc. Natl. Acad. Sci. USA* **94**, 11339–11344 (1997).
24. Kumamoto, C.A. & Francetiç, O. Highly selective binding of nascent polypeptides by an *Escherichia coli* chaperone protein *in vivo*. *J. Bacteriol.* **175**, 2184–2188 (1993).
25. Hartl, F.U., Lecker, S., Schiebel, E., Hendrick, J.P. & Wickner, W. The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* **63**, 269–279 (1990).
26. Park, S., Liu, G., Topping, T.B., Cover, W.H. & Randall, L.L. Modulation of folding pathways of exported proteins by the leader sequence. *Science* **239**, 1033–1035 (1988).
27. Fekkes, P., van der Does, C. & Driessen, A.J.M. The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J.* **16**, 6105–6113 (1997).
28. Xu, Z., Knafels, J.D. & Yoshino, K. Crystal structure of the bacterial protein export chaperone SecB. *Nature Struct. Biol.* **7**, 1172–1177 (2000).
29. Muren, E.M., Suciu, D., Topping, T.B., Kumamoto, C.A. & Randall, L.L. Mutational alterations in the homotetrameric chaperone SecB that implicate the structure as dimer of dimers. *J. Biol. Chem.* **274**, 19397–19402 (1999).
30. Knoblauch, N.T. *et al.* Substrate Specificity of the SecB Chaperone. *J. Biol. Chem.* **274**, 34219–34225 (1999).
31. Randall, L.L., Topping, T.B., Suciu, D. & Hardy, S.J. Calorimetric analyses of the interaction between SecB and its ligands. *Protein Sci.* **7**, 1195–1200 (1998).
32. Hardy, S.J. & Randall, L.L. A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. *Science* **251**, 439–443 (1991).
33. Topping, T.B. & Randall, L.L. Determination of the binding frame within a physiological ligand for the chaperone SecB. *Protein Sci.* **3**, 730–736 (1994).
34. Fekkes, P. *et al.* Preprotein transfer to the *Escherichia coli* translocase requires the cooperative binding of SecB and the signal sequence to SecA. *Mol. Microbiol.* **29**, 1179–1190 (1998).
35. Kawasaki, S., Mizushima, S. & Tokuda, H. Membrane vesicles containing overproduced SecY and SecE exhibit high translocation ATPase activity and countermovement of protons in a SecA- and presecretory protein-dependent manner. *J. Biol. Chem.* **268**, 8193–8198 (1993).
36. Joly, J.C. & Wickner, W. The SecA and SecY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids. *EMBO J.* **12**, 255–263 (1993).
37. van Voorst, F., van der Does, C., Brunner, J., Driessen, A.J.M. & de Kruijff, B. Translocase-bound SecA is largely shielded from the phospholipid acyl chains. *Biochemistry* **37**, 12261–12268 (1998).
38. Schiebel, E., Driessen, A.J.M., Hartl, F.U. & Wickner, W. $\Delta\mu H+$ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* **64**, 927–939 (1991).
39. van der Wolk, J.P., de Wit, J.G. & Driessen, A.J.M. The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events. *EMBO J.* **16**, 7297–7304 (1997).
40. Economou, A. & Wickner, W. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* **78**, 835–843 (1994).
41. Kaufmann, A., Manting, E.H., Veenendaal, A.K., Driessen, A.J.M. & van der Does, C. Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. *Biochemistry* **38**, 9115–9125 (1999).
42. Hanein, D. *et al.* Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* **87**, 721–732 (1996).
43. Beckmann, R. *et al.* Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science* **278**, 2123–2126 (1997).
44. Ménétret, J. *et al.* The structure of ribosome-channel complexes engaged in protein translocation. *Mol. Cell* **6**, 1219–1232 (2000).
45. Meyer, T.H. *et al.* The bacterial SecY/E translocation complex forms channel-like structures similar to those of the eukaryotic Sec61p complex. *J. Mol. Biol.* **285**, 1789–1800 (1999).
46. Hamman, B.D., Chen, J.C., Johnson, E.E. & Johnson, A.E. The aqueous pore through the translocon has a diameter of 40–60 Å during cotranslational protein translocation at the ER membrane. *Cell* **89**, 535–544 (1997).
47. Tani, K., Tokuda, H. & Mizushima, S. Translocation of ProOmpA possessing an intramolecular disulfide bridge into membrane vesicles of *Escherichia coli*. Effect of membrane energization. *J. Biol. Chem.* **265**, 17341–17347 (1990).
48. Manting, E.H., van der Does, C., Remigy, H., Engel, A. & Driessen, A.J.M. SecYEG assembles into a tetramer to form the active protein translocation channel. *EMBO J.* **19**, 852–861 (2000).
49. Karamanou, S. *et al.* A molecular switch in SecA protein couples ATP hydrolysis to protein translocation. *Mol. Microbiol.* **34**, 1133–1145 (1999).
50. Nakatogawa, H., Mori, H. & Ito, K. Two independent mechanisms down-regulate the intrinsic SecA ATPase activity. *J. Biol. Chem.* **275**, 33209–33212 (2000).
51. Eichler, J. & Wickner, W. Both an N-terminal 65-kDa domain and a C-terminal 30-kDa domain of SecA cycle into the membrane at SecYEG during translocation. *Proc. Natl. Acad. Sci. USA* **94**, 5574–5581 (1997).
52. Ramamurthy, V. & Oliver, D. Topology of the integral membrane form of *Escherichia coli* SecA protein reveals multiple periplasmically exposed regions and modulation by ATP binding. *J. Biol. Chem.* **272**, 23239–23246 (1997).
53. van der Does, C., Manting, E.H., Kaufmann, A., Lutz, M. & Driessen, A.J.M. Interaction between SecA and SecYEG in micellar solution and formation of the membrane-inserted state. *Biochemistry* **37**, 201–210 (1998).
54. Shilton, B. *et al. Escherichia coli* SecA shape and dimensions. *FEBS Lett.* **436**, 277–282 (1998).
55. Yahr, T.L. & Wickner, W.T. Evaluating the oligomeric state of SecYEG in preprotein translocase. *EMBO J.* **19**, 4393–4401 (2000).